

## Grammatical variation within Near-Standard German: connectors in corpora

ULRICH HERMANN WASSNER, MANNHEIM  
wassner@ids-mannheim.de

### Introduction

This article is, in a sense, a continuation of the one by Jacqueline Kubczak and Marek Konopka (this volume). Their comments on the framework that is the basis of our collective research are valid for this study as well. I am going to present a further example of our future project on grammatical variation in regional, medial, social, and especially text register varieties of Near-Standard German in quite a different area of linguistic research, namely, in the area of text linguistics. My special topic is a phenomenon we encountered in the course of our work on German connectors.

Connectors are much the same as conjunctions and sentence relating adverbs and particles of traditional grammar. The linguistic expressions linked by connectors are called 'connects'. The connect that immediately follows the connector or – in other cases – the one in which the connector is syntactically integrated is called 'internal'; the other connect we call 'external'. In example (1a), the connector is *obwohl*; the external connect is *Ich schreibe Mikrophon mit ph*; and the internal connect is *es ist jetzt auch mit f erlaubt*.

- (1) *Ich schreibe Mikrophon mit ph, obwohl es jetzt auch mit f erlaubt ist.*  
'I spell Mikrophon with *ph*, though nowadays it is allowed to spell it with *f* as well.'

Among the several syntactic subclasses of connectors, this article focuses on subordinating conjunctions. By *subordinating*, we mean that they require that the finite verb be in the sentence-final position in their internal connect, as is the case in (1); there the finite verb is underlined for easier identification. There are almost 100 subordinating conjunctions in German.<sup>1</sup>

What is striking about these kinds of connectors is that some of them have an alternative use we call paratactic. For connectors, this means that their internal connect is a 'normal main clause', an independent sentence.<sup>2</sup> It does not have

<sup>1</sup> For a list, see Pasch et al. (2003: 354, 418).

<sup>2</sup> Cf. Pasch et al. (2003: 369 ff and particularly 403 ff); [http://hypermedia.ids-mannheim.de/pls/public/sysgram.ansicht?v\\_id=1607](http://hypermedia.ids-mannheim.de/pls/public/sysgram.ansicht?v_id=1607); both with many recommendations for further reading; Zifonun et al. (1997 III: 2305f, and also I, 460 ff).

verb-last (VL) but instead, verb-second (V2) or verb-first position (V1) word order. With *obwohl*, this results in:

(2a) *Ich schreibe Mikrophon mit ph, obwohl: Es ist jetzt auch mit f erlaubt.*

(2b) *Ich schreibe Mikrophon mit ph, obwohl: Ist das jetzt nicht auch mit f erlaubt?*

Reference grammars and normative stylistics<sup>3</sup> admit only the variant (a) – VL; real texts show (b) – V2 and (c) – V1 as well. This observation is the starting point of our analysis.

Of the subordinating connectors with paratactic use, *weil* 'because' is the most prominent one and also the one most thoroughly studied. But in the course of this research, other connectors turned out to have the same feature; and as regards yet other ones, linguists have not yet come to any consensus whether they belong to this group. Moreover, explanations for this exceptional behavior are controversial and do not yet cover all cases. Here, work with large corpora might supply answers to yet unanswered and perhaps even – as regards traditional linguistic methods – unanswerable questions.

In our future project on grammatical variation in Near-Standard German, we will – using the huge electronic corpora of the Institut für Deutsche Sprache (IDS) – have to answer four questions concerning these connectors:

1. Exactly which German subordinating connectors actually occur paratactically, and which ones cannot?
2. What are the relative numbers of the occurrences of paratactic versus 'normal' subordinating use?
3. Can we find correlations between a certain use of a specific connector and the registers that we can identify in the corpora? In particular, is it true that the paratactic use of subordinating conjunctions occurs nearly exclusively in spoken<sup>4</sup> German?
4. Can theories be confirmed by corpus research or can even new explanations be found this way?

In this paper, I have to confine myself mainly to the first of these questions.

## 1 The range of subordinating connectors with paratactic use

Here I can only present some prototypical cases, highlighting methodological questions of corpus research and focusing on some heuristic hints that result from the corpus research so far.

<sup>3</sup> Cf. for *weil* Duden IX (2007: 996). For word order after *weil* or *obwohl* cf. Wahrig (2003: 493).

<sup>4</sup> Good empirical information about some of the relevant connectors in spoken German can be found in, e.g., Volodina (2007).

## 1.1 Uncontroversial cases serving as a litmus test: a heuristic and its problems

No one doubts – at least to the best of my knowledge – that the following connectors occur with V2 and V1: the four concessive ones *obwohl*, *obzwar*, *obschon*, and *obgleich*, the adversative version of *während*, causal *weil* and *zumal*, and *wobei* in its concessive meaning and use.<sup>5</sup> These belong to different semantic classes, which shows that specific semantic characteristics cannot generally be the reason for this behavior.

The corpora of the IDS deliver examples of paratactic use for these in sufficient numbers. Here are some clear-cut examples for *während*, *weil*, and *wobei*. For *obwohl*, see examples (2a–b). Notice that they all indeed sound very much like spoken German.

- (3) *Da hab ich gedacht um halb acht, die fangen da nämlich schon um halb acht an, während: hier in der Stadt fängt es immer erst um acht an.*  
'Then I thought at half past seven because there they start at half past seven, while here in town it starts only at eight.'
- (4) *Das ist keine Subvention, weil – es wird ja etwas gekauft.*  
'This is no subsidy because actually something is bought.'
- (5) *Deutsche Popmusik find ich einfach furchtbar. Wobei, manches davon ist ja schon fast wieder komisch.*  
'I hate German pop music. Some of it, however, is almost funny.'

Examples of this kind give us an interesting heuristic that can be used to find presumable instances of the paratactic use of one of these words in electronic corpora. Whenever the connector is followed by a colon (2a–b, 3), a dash (4), or a period (5), it is likely that its internal connect is an independent sentence, V2 or V1, instead of a subordinated one (VL). This fact, then, can help us find instances of appropriate cases for other connectors that have not yet been proven to be of the kind we are concerned with. (It is **not** a test to positively rule out that a given connector belongs to this class; it might have internal connects of the relevant kind without this graphic indicator preceding the connect.)

## 1.2 Example: *obgleich*

I want to exemplify the application of the test by using one of the undoubtedly paratactically used connectors that has not attracted much attention yet.

---

<sup>5</sup> For more information on these, see, e.g., Krutova (2005).

This case also demonstrates what problems arise in implementing the heuristic.<sup>6</sup>

### 1.2.1 *obgleich*,

In all the written texts found in the IDS-corpora (that is in the 'W-gesamt' corpus with its approximately 2,080,000,000 word forms), *obgleich* (whether capitalized or not) appears 12,220 times. In 31 of these examples, it is immediately followed by a comma – presumably indicative of a paratactic uses. But there is a caveat: from these, we have to subtract all those cases where the comma stems from the insertion of an additional subordinate clause (a relative or an adverbial clause) or a parenthetical. These factors do not signal V2 or V1 in the internal connect. An example of a 'bad' result for *obgleich* followed by comma is (6):

- (6) *Am 11. Mai 1802 schilderte Goethe Schiller diese Zwischenbilanz: „Mit meinem hiesigen Aufenthalt bin ich recht wohl zufrieden. Das Geschäft ist weiter gediehen, als ich hoffte, **obgleich**, wenn man strenge will, noch wenig geschehen ist. [...]“* (Frankfurter Rundschau, 28.8.1999, p. 37)  
 'On May 11th 1802 Goethe gave Schiller the following interim results: "I am rather content with my stay here. Business has been going better than I hoped although, strictly speaking, hardly anything has happened."'

Here we see that the internal connect of *obgleich* – *noch wenig geschehen ist* – still has VL. Cases of *obgleich* employed in metalinguistic enumerations (5b) should also be excluded:

- (7) *konzessiv (einräumend): obgleich, obwohl, (selten: ob, obschon, obzwar, wenn- gleich, wenn auch, wensschon, wiewohl, ungeachtet, gleichwohl) (Konjunktion (Wortart), in: Wikipedia - <http://de.wikipedia.org>, 2005)*

As far as I can see, all these cases can only be excluded by a human 'editor', based on his linguistic competence.

Thus, many findings have to be discarded. But fortunately, seven 'sound' results remain, cases of actual V2 internal connect. They are of the kind exemplified in (8):

- (8a) *Klubkollege Anton Ehmann wäre zu haben, **obgleich**, Manndecker besitzen die Innviertler selbst genug.* (Oberösterreichische Nachrichten, 31.7.1996, no p.)  
 '[His] teammate Anton Ehmann would be available. The "Innviertler" team, however, has enough defense players.'

<sup>6</sup> The examples come from the IDS-corpora. Thus, the empirical basis is somewhat broader than, e.g., in Dittmar – Bressemer (2005). The searching was performed with Cosmas II for the search term followed by a colon or a comma without a blank in between, or by a dash after a blank.



- (8b) „Ungeschmälerte Gehässigkeit“ hätten die über ihn ausgeschüttet, klagt Henning Voscherau, **obgleich**, Wunder nehmen täte ihn das nicht. (die tageszeitung, 18.9.2003, p. 22)  
 ‘They bullied him with “unrestricted spitefulness”, Voscherau complains, but that did not surprise him.’
- (8c) Der LASK hatte sogar noch großes Glück. „Und das zum ersten Mal überhaupt, seit ich in Linz bin“, wie Rausch bemerkte. **Obgleich**, belohnt wird bekanntlich der Tüchtige. Meistens jedenfalls. (Oberösterreichische Nachrichten, 18.11.1996, no p.)  
 ‘LASK even had great luck. “And this for the first time ever since I have been in Linz,” as Rausch remarked. Fortune favours the bold, however. At least most of the time.’

### 1.2.2 obgleich:

Our second indicator, the colon, introduces another problematic factor. There are 19 findings for *obgleich* followed by a colon, as exemplified by (7):

- (9a) Dass er schon seit fünf Jahren in Pension ist, glaubt man kaum – **obgleich**: In der Polizei hat sein Weggang eine Lücke gerissen. (Mannheimer Morgen, 16.8.2005, no p.)  
 ‘One does not believe that he has been retired for five years. He has, however, left a gap in the police department with his departure.’
- (9b) Wir rauchten nervös, **obgleich**: Was sollte geschehen? (die tageszeitung, 26.5.2001, p. III)  
 ‘We were smoking nervously. But ...: What would happen?’
- (9c) Fand den Artikel über Helge Mengel schockierend. (**Obgleich**: In einer Autogesellschaft kaum anders zu erwarten ...) (die tageszeitung, 29.5.1997, p. 24)  
 ‘I found the article about Helge Mengel shocking. (In a motor car society, however, you would hardly expect anything different ...)’

However, except for one single case, all these examples are of V2-type (9a–b). The exception, shown in (9c), again turns out to be problematic only after an evaluation by a human editor: here the internal connect is an ellipsis without a finite verb so that it is not obvious on the surface whether there should be VL or V2/V1.

### 1.2.3. obgleich –

The limitations of our test method are once more demonstrated by our third analysis. We searched for the sequence *obgleich* – blank – *dash* in the corpora. None of the 19 results contain a single relevant example. In 14 cases, the dash opens a parenthetical, and the internal connect ‘still’ has VL, shown in (10a–c). In one other case, the dash closes a parenthetical containing *obgleich* (10d) and in the remain-

ing four examples, there is another word between *obgleich* and the dash: a result that comes from the specific form of the query in Cosmas II, is shown in (10e):

- (10a) *Herr Scholl-Latour indes wird sich die militärische Operation [...] lieber live ansehen, weil er sowas interessant findet, **obgleich** – auch das gibt er zu bedenken – ein Kriegsschauplatz nie schön ist.* (die tageszeitung, 30.11.1990, p. 3)  
 'Mr. Scholl-Latour will be watching the military operation live because he is interested in such things although, he remarks, a theatre of war is never beautiful.'
- (10b) *Uneinsichtig blieb, warum der Stadtrat Energie und Zeit für die Prüfung eines Zusammengehens aufwendet, **obgleich** – wenn überhaupt – eigentlich nur die Schule profitieren kann.* (St. Galler Tagblatt, 3.9.1999, no p.)  
 'It remained unclear why the city council is spending time and effort for checking a fusion in advance although, if anyone at all, only the schools can profit from this.'
- (10c) **Obgleich** – oder gerade weil – die SVP von der FDP und allen Parteien links von ihr und erst recht von allen Grossmedien zum Hauptfeind Nummer 1 erkoren wurde, bedrängt sie im Wirtschaftskanton Zürich den Freisinn [...]. (Zürcher Tagesanzeiger, 20.3.1996, p. 25)  
 'Despite or just because the SVP has been made the enemy number 1 by FDP and all the parties left of it and all the more by the mass media, it is threatening the Freisinn party in canton Zürich.'
- (10d) *Lessing, wie Nietzsche, überblickte den Dichter, weil er ihn mit umfaßte, weil – oder **obgleich** – er unter anderem selbst einer war [...].* (T. Mann: Zu Lessings Gedächtnis, Erstv. 1929, in: Gesammelte Werke in zwölf Bänden mit einem Ergänzungsband, Bd. 10. Frankfurt a. M.: S. Fischer Verlag 1960, p. 252)  
 'Lessing like Nietzsche overlooked the poet because he embraced him, because – or despite – he also was one himself.'
- (10e) *Das Sterilisations-Gesetz von 1933 war, **obgleich** rassen- und geschlechtsneutral formuliert – es nannte primär psychische Sterilisationsgründe –, in seinen Auswirkungen nicht nur rassistisch, sondern auch spezifisch frauenfeindlich.* (Frankfurter Allgemeine, 1993, no p.)  
 'Though the sterilization law of 1933 was formulated in a race and gender neutral manner (it gave primarily psychological reasons for sterilization), it was not only racist in its effects but also specifically misogynist.'

In summary, these findings show us the scope – but also the limitations – of our heuristic, which we had derived from and applied to test cases that are clear from the beginning. However, not all cases are as straightforward as those mentioned above. Let us now apply the heuristic to controversial cases and – at first – to those not yet examined. Obviously, such cases are more interesting for our cor-

pus project. We can perhaps approach decisions concerning possible paratactic uses of subordinating connectors with our new-found heuristics.

### 1.3 Application of the test to hitherto unexplored cases

The list of clear cases in 1.1 contains candidates no one had thought of in the first place (in this branch of linguistic research) since everyone was focusing on *weil* alone. Especially *wobei* and the variants of *obwohl* have rarely or never been mentioned explicitly in previous research. Therefore, it is not entirely unlikely that other cases have been overlooked as well. Hence, in the end, we have to test each single item on the list of subordinating connectors (cf. the reference in Footnote 1) whether it can also be used paratactically, and we will do so using the heuristic described above.

#### 1.3.1 Theoretical predictions

There are subordinating connectors for which our theory predicts the impossibility of their use with a V2 or V1 internal connect. Grammis (the systematic internet grammar of the IDS) mentions two groups that should not be able to be used paratactically:<sup>7</sup>

1) those with a non-factive internal connect (that is hypothetical or conditional connectors in the broadest sense, including those that contain a semantic negation) such as *angenommen, dass; ausgenommen, dass; falls; gdw.; gesetzt, dass; gesetzten Fall, dass; im Fall; im Falle; insofern; insofern (...), als; insoweit; insoweit (...), als; so; sofern; unterstellt, dass; vorausgesetzt, dass; wenn; and*

2) those that cannot refer (as a confirmation or contradiction) to speech acts or propositional attitudes as well as to expectations resulting from them (that is, temporal connectors above all), such as *als; bevor; bis; bis dass; derweil; derweilen; ehe; indes; indessen; kaum dass; nachdem; nun; seit; seitdem; sintemal; sintemalen; sobald; solange; solange; sooft; sowie; währenddessen.*

Do these predictions prove to be correct?

#### 1.3.2. The case of *alldieweil*

An especially interesting case could be, for example, *alldieweil*, which has two basic meanings: a temporal one (resembling the English 'while, in the meantime, meanwhile') and a causal one (like *weil* 'because').<sup>8</sup> The question is, does *alldieweil* show a different behavior regarding the position of the verb in the internal connect according to its particular meaning?

Unfortunately (but interestingly), *alldieweil* is one of the rarer connectors. Even

<sup>7</sup> [http://hypermedia.ids-mannheim.de/pls/public/sysgram.ansicht?v\\_id=1607](http://hypermedia.ids-mannheim.de/pls/public/sysgram.ansicht?v_id=1607)

<sup>8</sup> This reflects the situation captured in dictionaries. Our research on German connectors has yielded the result that – just like *while* or German *während* – *alldieweil* also has another, adversative meaning that has to be taken into consideration.

in the more than two billion word forms of the corpora of the IDS, it only occurs 91 times. And so it is no wonder that there is no instance at all of  $A/_a$  *alldieweil* followed by a colon in W-gesamt. The situation is better for  $A/_a$  *alldieweil* followed by a comma: at least two instances were found. And indeed one of them – (11)<sup>9</sup> –

- (11) *Es stellt sich mir/uns die Frage, ob wir nicht allzusehr mit ökologischen Themen hantieren, für die es hier im Osten eben keine vorhergelaufene 20jährige Gesellschaftsdebatte aufzuweisen gibt. Das gilt es nachzuholen [...]. Alldieweil, dem muß eine Debatte zwischen Westgrünen und Ostbürgerrechtlern vorausgehen, aus welcher dann erst bundesrepublikanische Grundnenner auch praktikabel für B'90/Grüne hervorgehen, West wie Ost.* (die tageszeitung, 10.11.1994, p. 26)  
 'The question is whether we are not too heavily focussing on ecological issues for which there has not been a twenty-year debate in the Eastern society before. We have to make up for this. Because this has to be preceded by a debate between western greens and eastern civil rights activists, which will lead to a common ground for B'90/Grüne in the whole FRG, west as well as east.'

proves that *alldieweil* is also one of the subordinating conjunctions that can be used paratactically. As there are no further examples, we cannot say anything about a possible contrast to the other meanings. Incidentally, though, the second instance, shown in (12), allows us to address another problem of our test heuristic.

- (12) *Die vier angehenden Verkäuferinnen lernten alldieweil, belegte Brötchen so appetitlich zu garnieren, daß die Kundschaft darauf fliegt.* (Mannheimer Morgen, 2.9.1998, no p.)  
 'In the meanwhile, the four sales women in training were learning how to arrange sandwiches so deliciously that the customers like them.'

Some connectors are syntactically ambiguous; that is, they belong to different syntactic subclasses of connectors. So does *alldieweil* – it can be not only a conjunction but also an adverb<sup>10</sup>. These cases have to be excluded, ideally by a computer but if necessary, by a human interpreter. This should be easy as long as *alldieweil* alone is the prefield, since the following verb must then be finite. The middle field position, as in example (10), is more problematic for automatic exclusion. But to be sure, these cases are definitely not included in the topic of paratactic use of subordinating conjunctions.

It should be mentioned briefly that our third test,  $A/_a$  *alldieweil* with a dash, produced a single occurrence in W-gesamt for which the same remarks apply as to *alldieweil* with comma.

<sup>9</sup> Where, by the way, *alldieweil* indisputably has the adversative meaning ('however').

<sup>10</sup> To be exact: an adverb that can by itself be the prefield or that can stand somewhere in the middle field of a sentence (but nowhere else).

## 1.4 Attempt of an application of the test to a controversial case: *damit*

In a recent talk, “Die Verletzung des verbalen Rahmens am Beispiel der kausalen Subjunkturen *weil*, *obwohl*, *damit* und *wobei*”, Růžena Kozmová lists *damit* as a paratactically usable subjunctor. I tend to challenge this assumption. At any rate, it should clearly be tested. But, while our heuristic is easily applicable to *alldieweil* not least because of the small number of examples, which makes it easy for the human researcher to handle them where the computer fails, *damit* occurs more than one million times in the corpora (to be exact: 1,490,523). And even *damit* with a colon is too frequent to be dealt with, and the corpus search for it results in a great amount of garbage, not least because *damit* often occurs as a pronominal adverb ‘with that, with it’ – such cases should be excluded first.<sup>11</sup> The search for *damit* plus a colon yields 1,206 examples, most of them in formulations not relevant for us, such as *Was meinte er damit: ...*, *Nicht genug damit: ....* I cannot yet see any way to handle such a large amount of data in a reasonable time span. The combination of *damit* followed by a dash has as many as 2,339 occurrences; in many of these, the dash stems from a subsequent parenthetical. And the findings for *damit* with a comma are even less feasible – the search yields 69,919 examples, as often appropriate as inappropriate.

## 1.5 Remnants

For the large group of the remaining subordinating connectors, no research concerning possible paratactic use is known and no theoretical assumptions have been proposed yet to explain one behavior or the other. These cases have to be taken up in future research.

As our project is only beginning, we do not have concrete answers to the other three questions. At present, we can only show the kind of specific inquiries connected with these questions.

## 2 The relative frequencies of the occurrences of paratactic and standard subordinating use

The way to explore paratactic uses of subordinating connectors that immediately comes to mind is to examine the position of the finite verb: does it occur

<sup>11</sup> To turn to subcorpora instead is no solution. To take another example: *wobei* appears 1,886 times in TAGGED (see below). As with *damit*, we have the problem of homonymous uses; often it cannot even be determined whether *wobei* is meant as a connector or as a relative adverb – in both cases it has VL. All this indicates that an automatic evaluation would create great problems; at the same time, the numbers are much too large for a manual evaluation.

at the end of a clause or a sentence, or not? For an investigation of this kind, we need morpho-syntactically annotated corpora. The IDS-corpora contain a subset of this kind, called TAGGED, comprising over 26 million word forms. Tagged corpora should give us new possibilities to search for internal connects with V2 versus those with VL. But, unfortunately, the current state of TAGGED is not very helpful: there are many tagging errors and other problems that hamper their use. Thus the easy idea of searching for a finite verb followed by a full stop, question, or exclamation mark is not applicable: we cannot receive all and only finite verbs when searching for them.

The result of the corpus research should be matrices that show, for example, the number of occurrences of *obwohl*: *obgleich*: *obschon*: *obzwar* with VL or with V2/V1 in register 1 to n. But at the moment, this partial corpus only shows things such as the fact that the four variants of concessive *ob- plus x* themselves are by no means equally frequent but that they are noticeably graded: *obwohl* 3,962 findings, *obgleich* 205, *obschon* 63, *obzwar* 4 (of which, incidentally, only one has a finite verb, which is VL).<sup>12</sup> The significant differences in number are at least one first quantitative result that might have implications concerning the relevance of the variants. It is common knowledge that *obwohl* is the unmarked and most frequent form of the four. But the other three seem intuitively to be of equal weight so that a quantitative relation of about, say, 100 : 1 : 1 : 1 could be expected. In reality the three *obwohl*-siblings show very different frequencies somewhere in the region of three- : two- : one-digit.

The questions remain: What is the reason for this? Is there a motivation at all? Could register be the reason? Or is this a case of outdated usage? A treatment of the latter question would require the possibility of analysing results diachronically.

### 3 Variants in varieties

The third question for future research is: can we find correlations of a distinct use of a certain connector with certain registers that we can distinguish in the corpora? With appropriate corpora, we should be able to deal with issues not only concerning German standard language as a whole but also variations across different varieties of this language, be they synchronic or diachronic. Here are some theses:

#### Dimension 1: spoken versus written language

This dimension is especially relevant. Most researchers consider *weil* + V2 an exclusive feature of spoken German. This raises the question: do corpora of written

<sup>12</sup> Of course, one cannot make any, let alone statistically significant, statements (which we should be aiming for) on the basis of such small numbers. Such cases are at best impressionistically illustrative.

texts contain no instances of this phenomenon at all? On the other hand, spoken language is filled with *anacolutha* etc. This fact casts serious doubt as to whether the putative instances of V2 usage observed in the corpus data are indeed cases of genuine V2 usage – they might be cases of performance errors caused by grammatically irrelevant factors. Spoken utterances are sometimes hardly comprehensible, let alone grammatically analyzable. This is a severe problem for any corpus-based research.

### **Dimension 2: diachronic variants, linguistic change**

The phenomenon is generally taken to be a recent one. But we find examples of *wiewohl* and *obgleich* plus V2 at least as early as 1880 (cf. Paul 1970: 298). In this domain, substantial diachronic observations could be made by comparing sufficiently large corpora containing texts from different times.

### **Dimension 3: regional variants, dialects**

Many researchers postulated *weil* + V2 to occur exclusively in Southern German, thus being a regional variant only of some dialects. Corpus research has provided evidence that this is not the case. Nevertheless, adequate corpora may show that *weil* + V2 is widespread in spoken colloquial language mainly in that region. Here it occupies the positions of and has the functions taken by *denn* in the northern German area, which again is restricted largely to written language in the south. Thus, corpora should not only contain written as well as spoken texts but also a sample of a large variety with respect to their regional and temporal origin, each in ample size.

### **Dimension 4: sociolect, register**

For the research concerning registers like academic prose or newspaper language, we will have to define 'virtual subcorpora', a future possibility that Cosmas II offers.

Lastly, adequate corpora should enable us to correlate all those dimensions with each other.

## **4 Finding explanations outside of corpora**

A further (but no less important) issue to be addressed is whether working with corpora may yield preliminary explanations for the observed facts and whether it allows us to test current explanations of our phenomenon. In particular, are theoretical assumptions confirmed by the results of corpus research or can we, perhaps, even find new explanations this way?

Here we leave the field of grammar in the narrower sense and step into semantics, which usually raises great challenges for the work with corpora using

the traditional tools; for example, determining correlations between one of the *weil*-variants and the reference to one of Sweetser's (1990) three 'domains' will presumably require a good deal of human interpreting.

Also, corpus findings may yield hints for the evaluation of certain uses. Are they errors, i.e., are they (grammatically, stylistically) poor language? Or are the norms inappropriate? A normative characterization as 'a deviation from the standard norm' may itself be mistaken; we might be observing genuine cases of language change – newly emerging forms with their own function.

But, to reiterate again, all such considerations are a long way off and all these theses are the objects of future research. At the moment, we can at best keep pursuing the phase of gathering material.

## References

- DITTMAR, N. – BRESSEM, J. (2005): Syntax, Semantik und Pragmatik des kausalen Konnektors *weil* im Berliner 'Wendekorpus' der neunziger Jahre. In: Schwitalla, J. – Wegstein, W. (eds.), *Korpuslinguistik deutsch: synchron – diachron – kontrastiv*. Würzburger Kolloquium 2003, Tübingen: Niemeyer, 99–124.
- DUDENREDAKTION (ed.) (2007): *Duden Band 9. Richtiges und gutes Deutsch. Wörterbuch der sprachlichen Zweifelsfälle*. 6., vollst. überarb. Aufl., Mannheim u.a.: Dudenverlag.
- KRUTOVA, M. (2005): *Subordinatoren des Deutschen als parataktische Konnektoren*. Unpublished Masters Thesis, Heidelberg.
- PASCH, R. – BRAUSSE, U. – BREINDL, E. – WASSNER, U. H. (2003): *Handbuch der deutschen Konnektoren. Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfen (Konjunktionen, Satzadverbien und Partikeln)*. Berlin/New York: de Gruyter.
- PAUL, H. (1970): *Prinzipien der Sprachgeschichte*. Studienausgabe der 8. Aufl. 1968, Tübingen: Niemeyer. 1. Aufl. 1880.
- SWEETSER, E. E. (1990): *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- VOLODINA, A. (2007): *Konditionale und kausale Relationen im gesprochenen Deutsch*. Unpublished Dissertation, Heidelberg.
- WAHRIG (2003): *Fehlerfreies und gutes Deutsch*. Gütersloh/München: Wissen Media Verlag.
- ZIFONUN, G. – HOFFMANN, L. – STRECKER, B. et al. (1997): *Grammatik der deutschen Sprache*. 3 Bände. Berlin/New York: de Gruyter.